



SPARSE2DGS: SPARSE-VIEW SURFACE RECONSTRUCTION USING 2D GAUSSIAN SPLATTING WITH DENSE POINT CLOUD

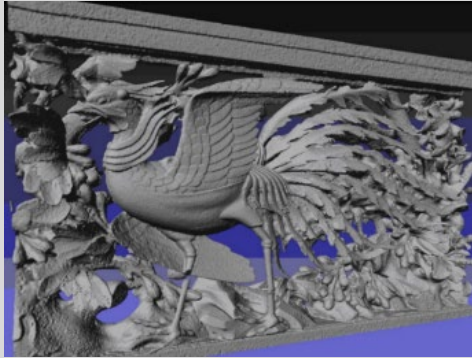
Natsuki Takama¹, Shintaro Ito¹, Koichi Ito¹, Hwann-Tzong Chen², and Takafumi Aoki¹

¹ Graduate School of Information Sciences, Tohoku University, Japan

² Department of Computer Science, National Tsing Hua University, Taiwan

IEEE ICIP 2025

Multi-View 3D Reconstruction



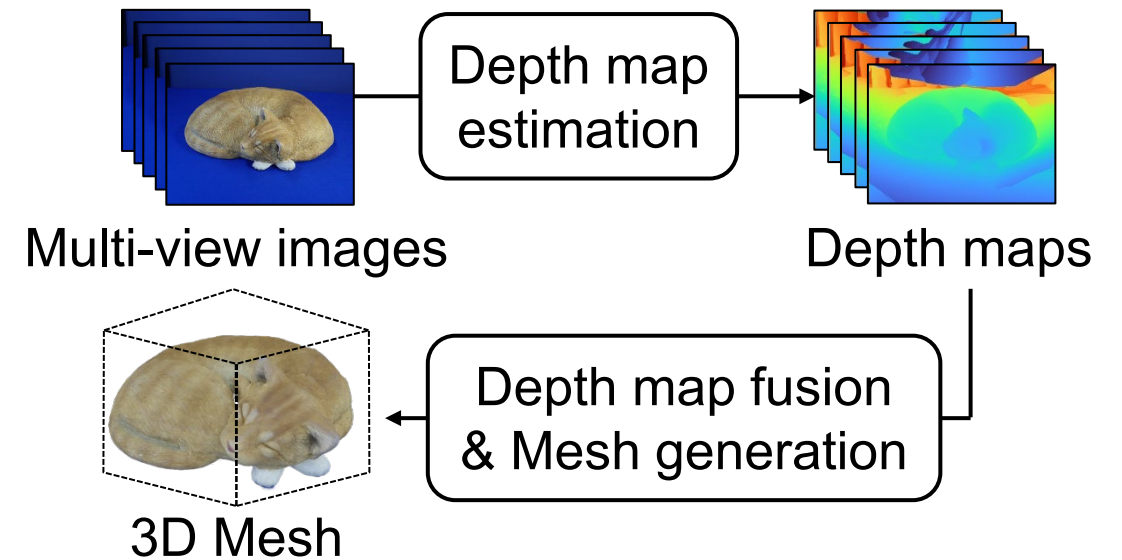
Digital archives [1]



Free viewpoint video [2]

Applications of 3D data

- Reconstruct the 3D shape of a target object from images taken at multiple viewpoints
- Estimate a depth map for each viewpoint and then integrate them to reconstruct the 3D shape of the target

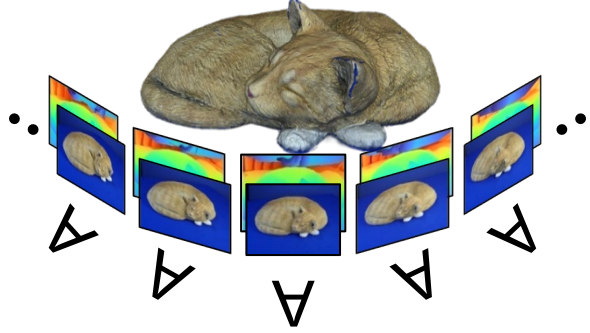
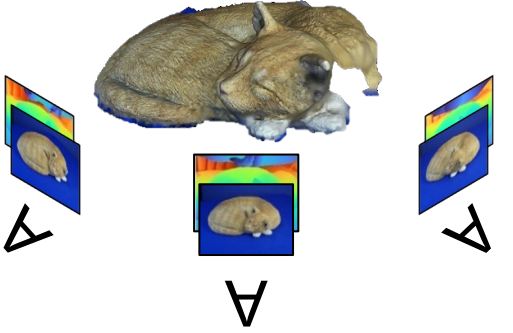


Pipeline of multi-view 3D reconstruction

[1] Sakai et al., "Phase-based window matching with geometric correction for multi-view stereo," IEICE Trans. Fundamentals, 2015.

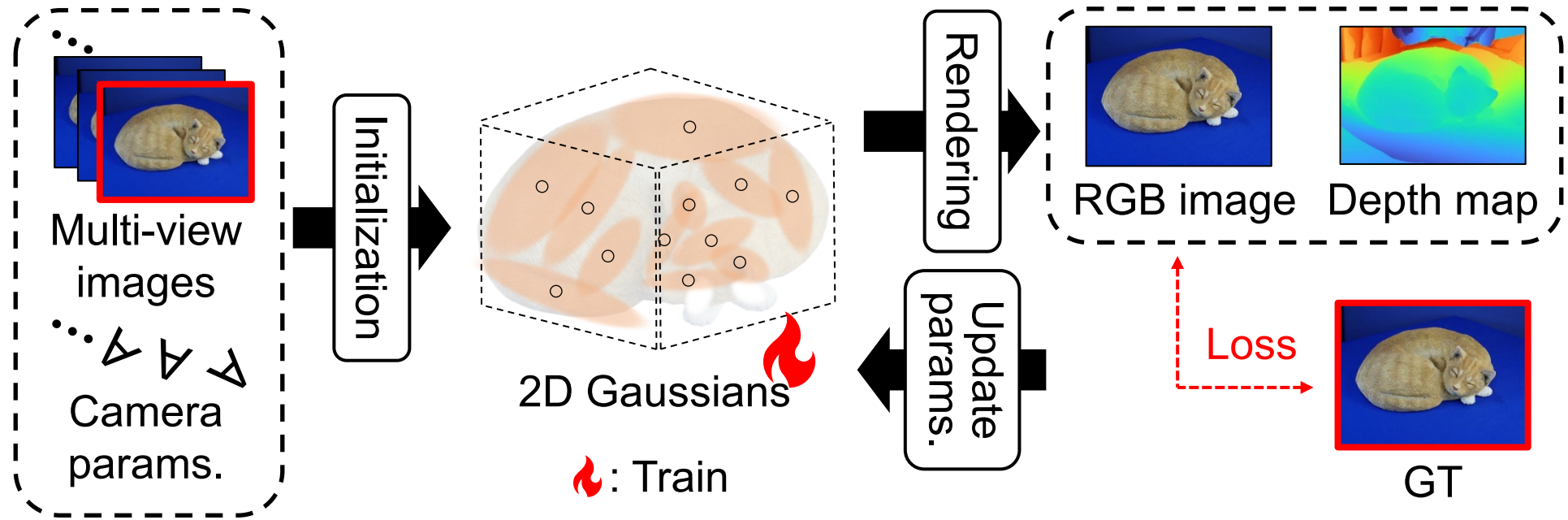
[2] <https://global.canon/ja/vvs/about/>

Multi View vs. Sparse View

	Multi-view 3D reconstruction	Sparse-view 3D reconstruction
Overview		
Capturing time	Long	Short
Scene	Limited	Widely applicable
Accuracy	High	Low

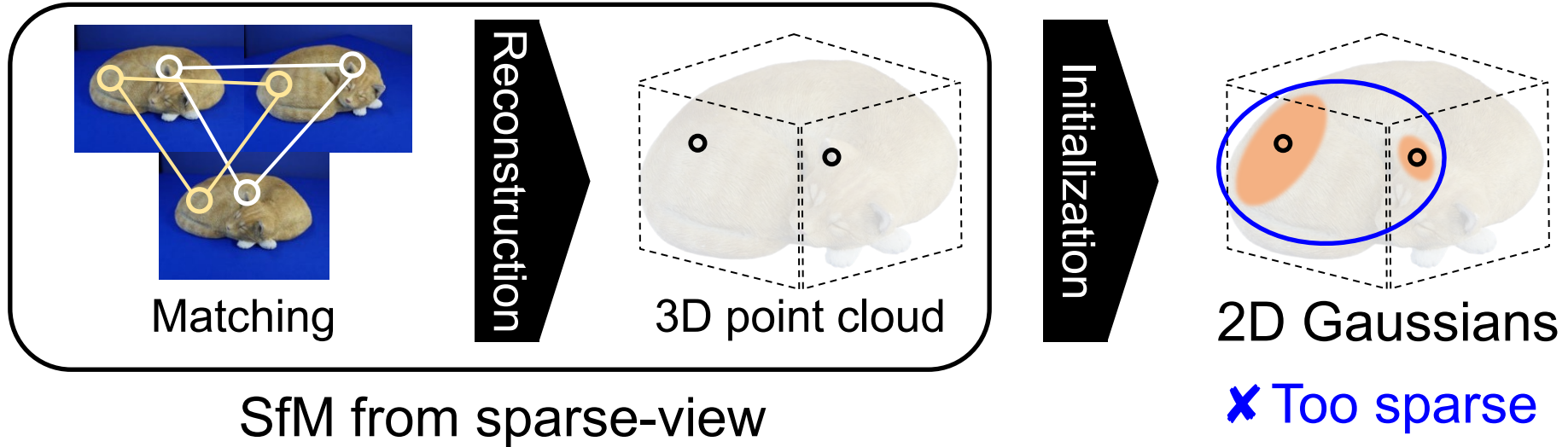
- Need to improve the accuracy of sparse-view 3D reconstruction
- Adapt 2DGS [3], a fast and high-accuracy multi-view surface reconstruction method, for a sparse-view scenario

2D Gaussian Splatting (2DGS) [3]



- Represent a radiance field by a set of 2D Gaussians, which are ellipses defined by parameters (e.g., position, color, opacity)
- Reconstruct a high-fidelity mesh using depth maps rendered from a geometrically accurate radiance field
- Require an accurate 3D point cloud of the entire scene to initialize 2D Gaussians

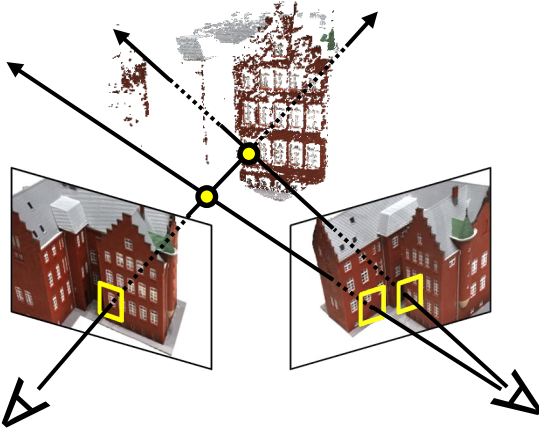
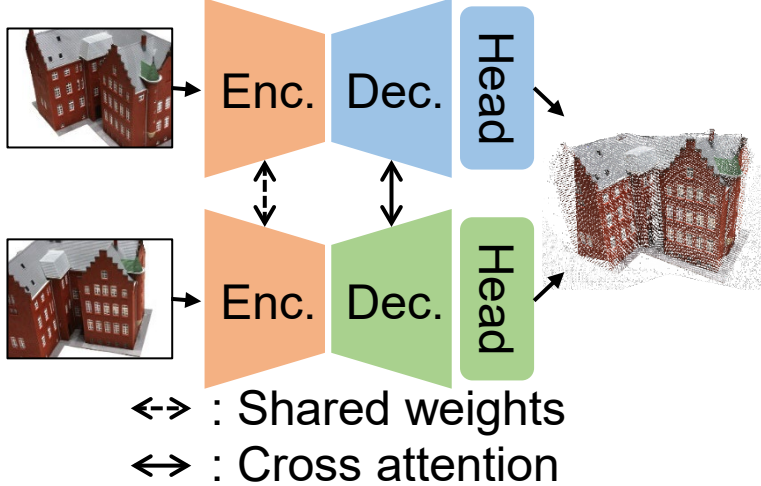
2DGS for Sparse-View Scenario



- 2D Gaussians are initialized from an accurate point cloud obtained from Structure from Motion (SfM)
- The 3D point cloud reconstructed from sparse-view images is too sparse
- An insufficient number of initial Gaussians results in significantly decreasing the accuracy of surface reconstruction in 2DGS

Leverage MVS and DUST3R to acquire a dense point cloud for initialization

MVS and DUS3R in Sparse-View Scenario

	COLMAP MVS [4]	DUS3R [5]
Overview		
Approach	Image matching	Neural network
Accuracy	High	Middle
# of 3D points	Small	Large

Combine MVS and DUS3R for obtaining an accurate and dense 3D point cloud

[4] J.L. Schönberger et al., "Pixelwise view selection for unstructured multi-view stereo," ECCV, 2016.

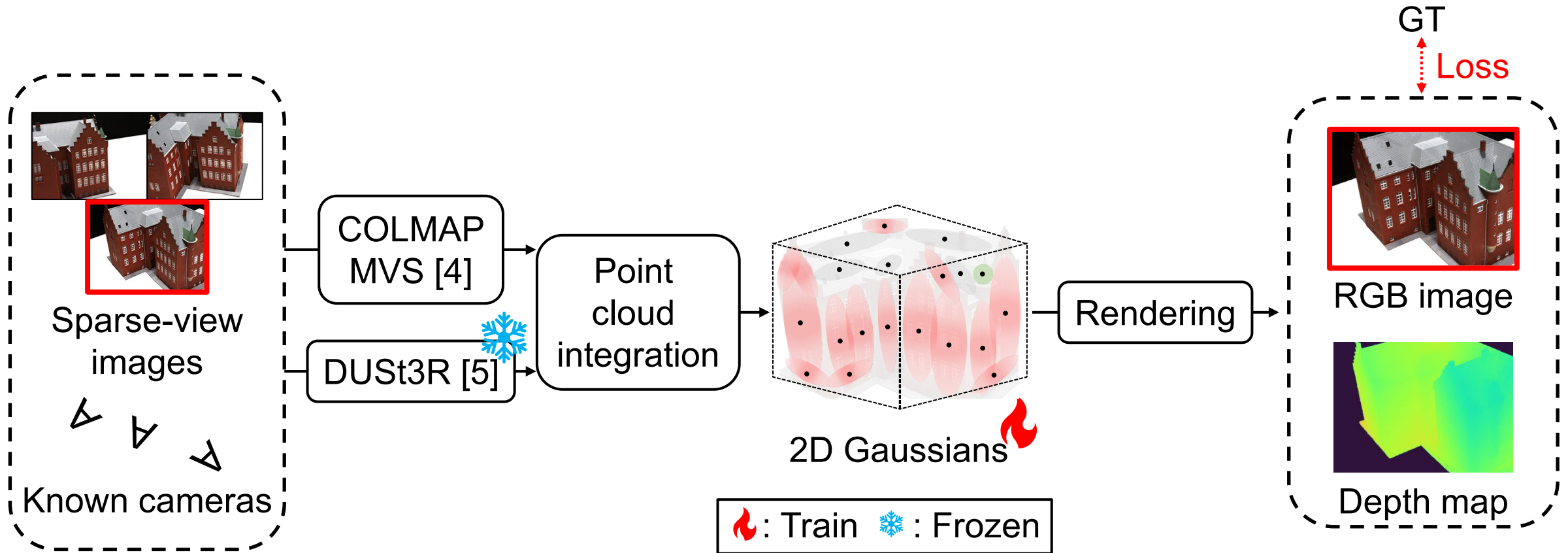
[5] S. Wang et al., "DUS3R: Geometric 3D vision made easy," CVPR, 2024.

Objective

Accurate 3D reconstruction from 3 views

- Improve the accuracy of sparse-view 3D reconstruction using 2DGS by initializing 2D Gaussians with a dense 3D point cloud
- Combine 3D point clouds by COLMAP MVS and DUS3R to obtain an accurate and dense point cloud for 2D Gaussian initialization
- Demonstrate the effectiveness of the proposed method through experiments on DTU dataset, where 3 views are input as a sparse-view scenario

Proposed Method

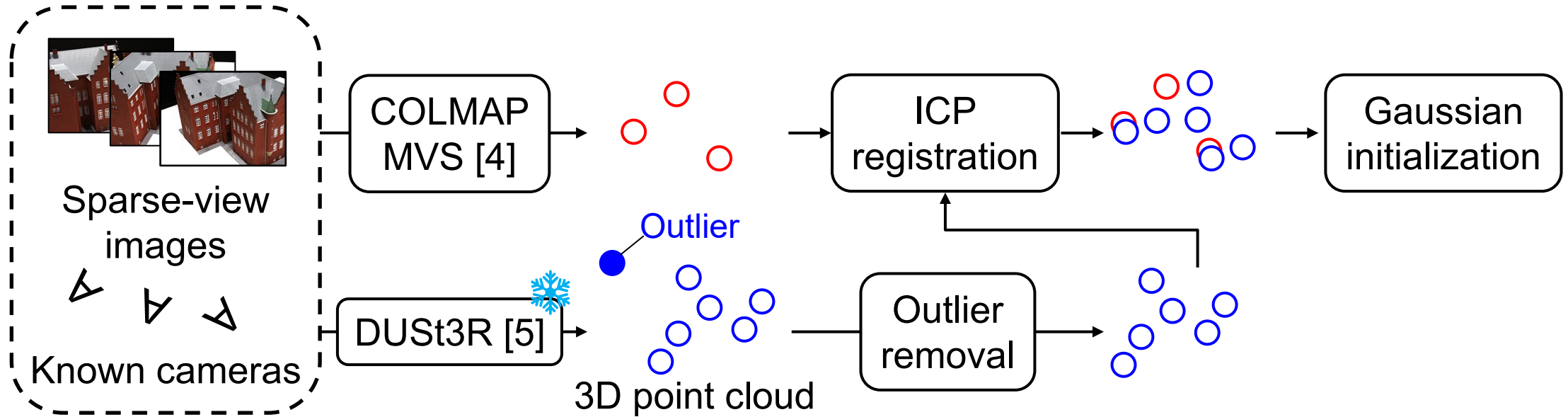


- Improve the accuracy of sparse-view surface reconstruction using 2DGS by initializing 2D Gaussians with a point cloud integrated from the outputs of COLMAP MVS [4] and DUST3R [5]

[4] J.L. Schönberger et al., "Pixelwise view selection for unstructured multi-view stereo," ECCV, 2016.

[5] S. Wang et al., "DUST3R: Geometric 3D vision made easy," CVPR, 2024.

Initialization Using COLMAP MVS [4] and DUS3R [5]



- Remove outliers from the point cloud reconstructed by DUS3R
- Align the point clouds of COLMAP MVS and DUS3R using the ICP algorithm
- Use the integrated point cloud as the initial positions of 2D Gaussians

[4] J.L. Schönberger et al., "Pixelwise view selection for unstructured multi-view stereo," ECCV, 2016.

[5] S. Wang et al., "DUS3R: Geometric 3D vision made easy," CVPR, 2024.

Experiments

- Evaluate the accuracy of a mesh model reconstructed by proposed method and compare the accuracy with other methods
- Input 3 images and their corresponding camera parameters
 - A standard configuration in sparse-view scenarios
- Reconstruct a mesh model by integrating depth maps using Truncated Signed Distance Fusion (TSDF) [9]
- Compared methods
 - 2DGS [3], COLMAP MVS [4], and DUS3R [5]
 - NeRF-based standard methods for sparse-view scenarios
 - SparseNeuS [6], ReTR [7], and UFORecon [8]

[3] B. Huang et al., “2D gaussian splatting for geometrically accurate radiance fields,” ACM Trans. Graph, 2024.

[4] J.L. Schönberger et al., “Pixelwise view selection for unstructured multi-view stereo,” ECCV, 2016.

[5] S. Wang et al., “DUS3R: Geometric 3D vision made easy,” CVPR, 2024.

[6] X. Long et al., “SparseNeuS: Fast generalizable neural surface reconstruction from sparse views,” ECCV, 2022.

[7] Y. Liang et al., “ReTR: Modeling rendering via transformer for generalizable neural surface reconstruction,” NeurIPS, 2023.

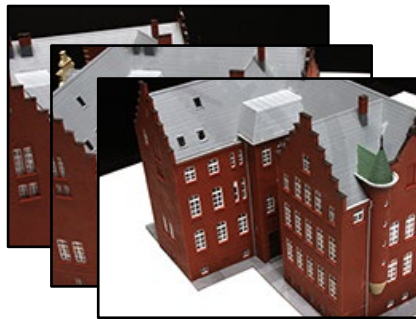
[8] Y. Na et al., “UFORecon: Generalizable sparse-view surface reconstruction from arbitrary and unfavorable sets,” CVPR, 2024.

[9] Q. Zhou et al., “Open3D: A Modern Library for 3D Data Processing,” arXiv, 2018

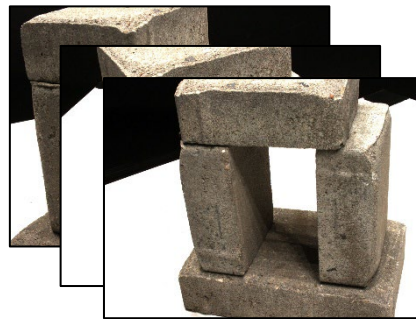
Dataset and Metric

■ Dataset : DTU dataset [10]

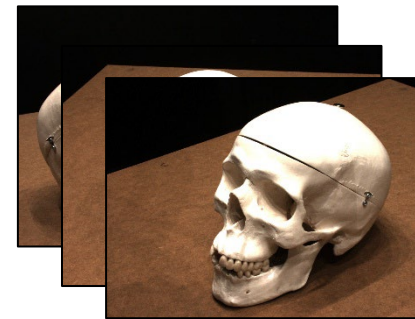
- Consist of multi-view RGB images, camera parameters, ground truth 3D point clouds
- Provide 3 images with little overlap and their corresponding camera parameters as input



scan24



scan40



scan65

■ Metric : Chamfer Distance (CD) ↓

- Calculate the distance between the reconstructed mesh and the ground-truth point cloud

Quantitative Results

Method	Scan ID												Avg.
	24	37	40	55	63	65	83	105	106	114	118	122	
SparseNeuS [6]	4.28	4.76	3.69	1.78	2.93	2.69	1.99	1.89	1.93	1.12	2.24	1.88	2.60
ReTR [7]	3.37	3.55	3.43	2.90	2.87	3.05	2.33	2.04	2.79	1.52	2.34	2.06	2.69
UFORecon [8]	1.51	2.61	1.93	1.47	1.58	1.80	1.54	1.34	1.20	0.65	1.26	1.25	1.51
COLMAP [5]	2.29	3.29	1.64	2.03	2.33	4.51	4.48	4.02	2.47	1.87	2.52	1.80	2.77
2DGS [3]	---	---	5.31	1.60	3.57	2.29	---	2.73	2.20	1.36	2.20	1.63	2.54
DUST3R [4]	1.34	3.28	1.76	2.00	2.57	2.21	1.85	1.78	2.52	1.38	1.92	2.45	2.09
Ours	1.01	2.66	1.56	1.36	1.83	1.63	1.91	1.33	1.53	0.64	1.29	1.34	1.51

Best results are highlighted as 1st, 2nd, and 3rd, respectively.

“---” means a failure case.

■ In most scans, our method shows 1st or 2nd reconstruction accuracy

[3] B. Huang et al., “2D gaussian splatting for geometrically accurate radiance fields,” ACM Trans. Graph, 2024.

[4] J.L. Schönberger et al., “Pixelwise view selection for unstructured multi-view stereo,” ECCV, 2016.

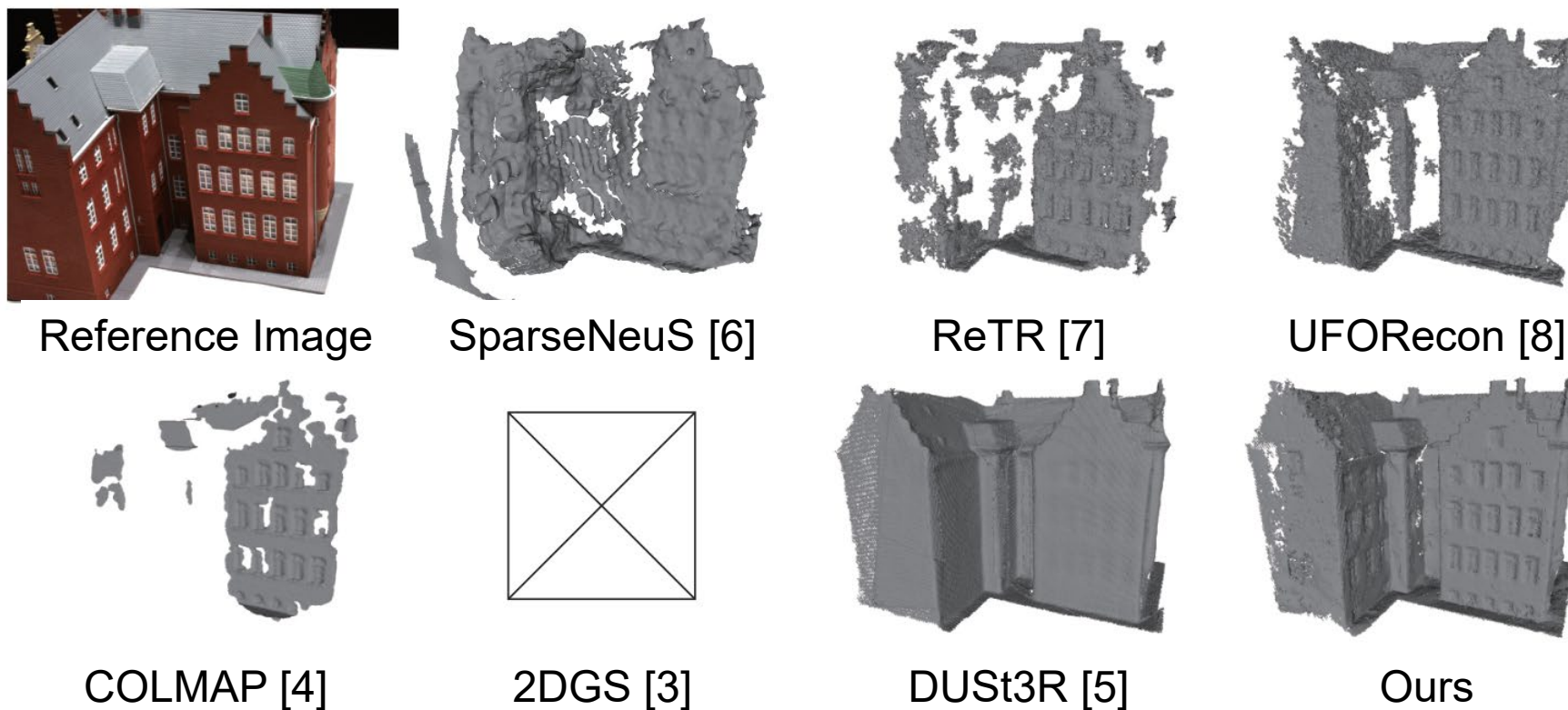
[5] S. Wang et al., “DUST3R: Geometric 3D vision made easy,” CVPR, 2024.

[6] X. Long et al., “SparseNeuS: Fast generalizable neural surface reconstruction from sparse views,” ECCV, 2022.

[7] Y. Liang et al., “ReTR: Modeling rendering via transformer for generalizable neural surface reconstruction,” NeurIPS, 2023.

[8] Y. Na et al., “UFORecon: Generalizable sparse-view surface reconstruction from arbitrary and unfavorable sets,” CVPR, 2024.

Qualitative Results



- Our method can reconstruct a wide area with high accuracy

[3] B. Huang et al., “2D gaussian splatting for geometrically accurate radiance fields,” ACM Trans. Graph, 2024.

[4] J.L. Schönberger et al., “Pixelwise view selection for unstructured multi-view stereo,” ECCV, 2016.

[5] S. Wang et al., “DUST3R: Geometric 3D vision made easy,” CVPR, 2024.

[6] X. Long et al., “SparseNeuS: Fast generalizable neural surface reconstruction from sparse views,” ECCV, 2022.

[7] Y. Liang et al., “ReTR: Modeling rendering via transformer for generalizable neural surface reconstruction,” NeurIPS, 2023.

[8] Y. Na et al., “UFORecon: Generalizable sparse-view surface reconstruction from arbitrary and unfavorable sets,” CVPR, 2024.

Conclusion and Future Work

■ Conclusion

- Proposed a method to improve the accuracy of sparse-view 3D reconstruction using 2DGS with a dense point cloud
- Demonstrated the effectiveness of the proposed method through the experiments

■ Future Work

- Explore enhancing the accuracy of sparse-view 3D reconstruction by introducing a reprojection-based loss and consider the use of NVS

There are some supplemental materials (Ablation, Why 2DGS not 3DGS, ...).
Please let me know if you have any questions. Thank you for your attention!

Supplementals

- Here are the supplementary materials. Please let me know if you have any questions.
 - Ablation study of initialization
 - Why 2DGS instead of 3DGS ?
 - Reconstruction results of other scenes
 - What is DUS_t3R ?

Ablation Study of Initialization (Quantitative)

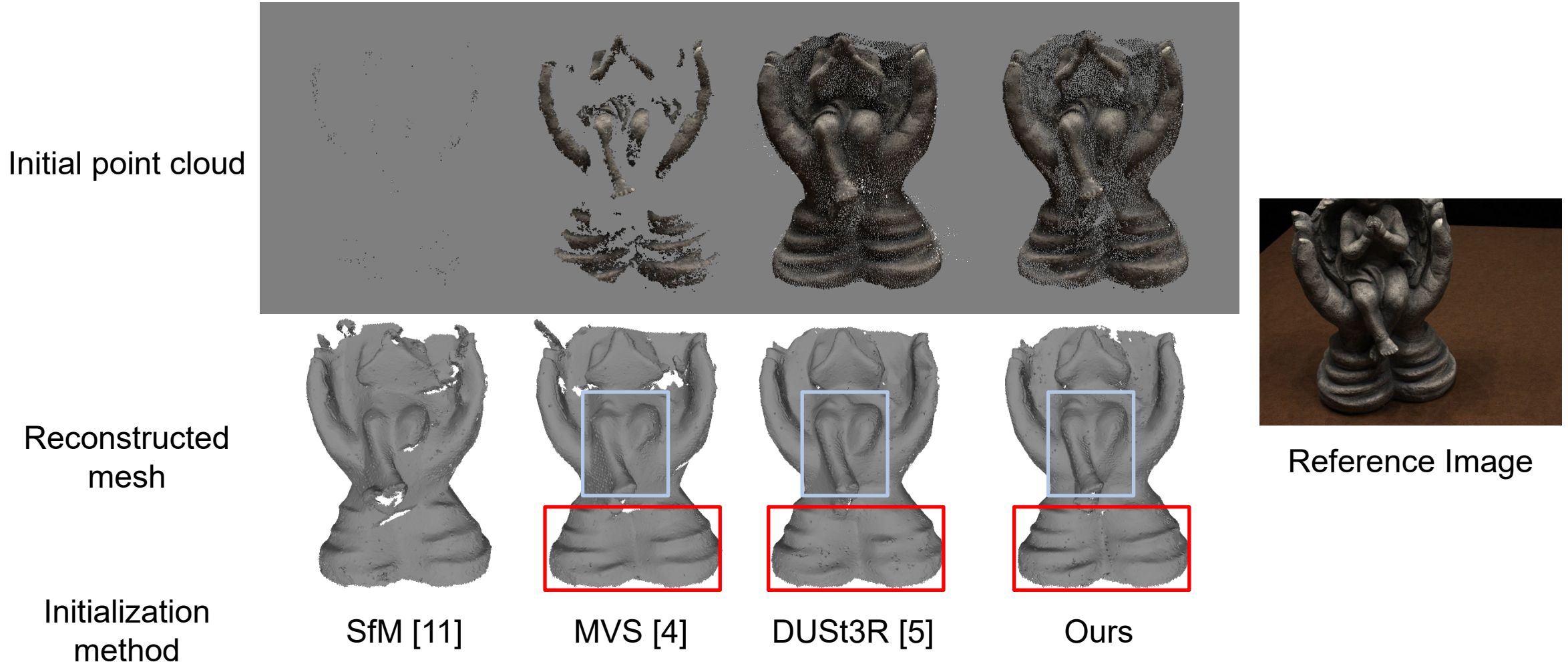
Init. method of 2DGS	Scan ID												Avg.
	24	37	40	55	63	65	83	105	106	114	118	122	
SfM [11]	---	---	5.31	1.60	3.57	2.29	---	2.73	2.20	1.36	2.20	1.63	2.54
MVS [4]	2.87	3.25	2.50	1.05	1.84	2.36	1.78	1.42	2.00	0.65	1.76	1.31	1.90
DUST3R [5]	1.03	3.00	1.90	1.39	2.47	1.67	1.89	1.23	1.97	0.99	1.40	2.21	1.76
Ours	1.01	2.66	1.56	1.36	1.83	1.63	1.91	1.33	1.53	0.64	1.29	1.34	1.51

Best results are highlighted as 1st, 2nd, and 3rd, respectively.
“---” means a failure case.

- We initialize with each method and compare the reconstruction accuracy based on 2DGS
- Our proposed initialization method demonstrates better reconstruction accuracy

[4] J.L. Schönberger et al., “Pixelwise view selection for unstructured multi-view stereo,” ECCV, 2016.
[5] S. Wang et al., “DUST3R: Geometric 3D vision made easy,” CVPR, 2024.
[11] J.L. Schönberger et al., “Structure-from-Motion Revisited,” CVPR, 2016.

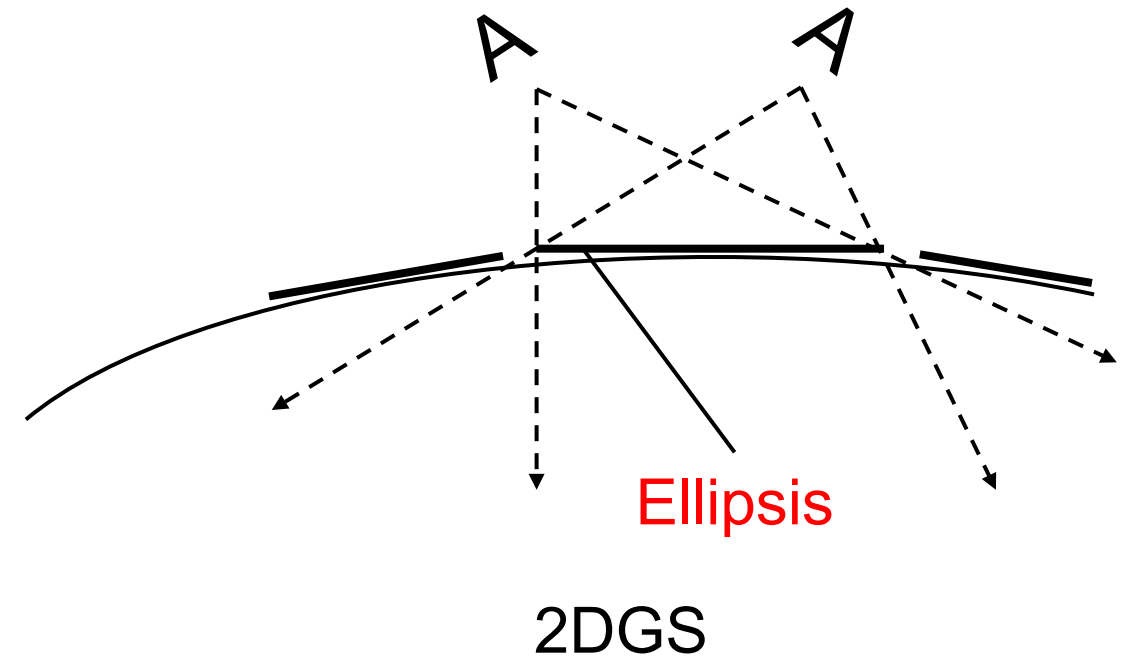
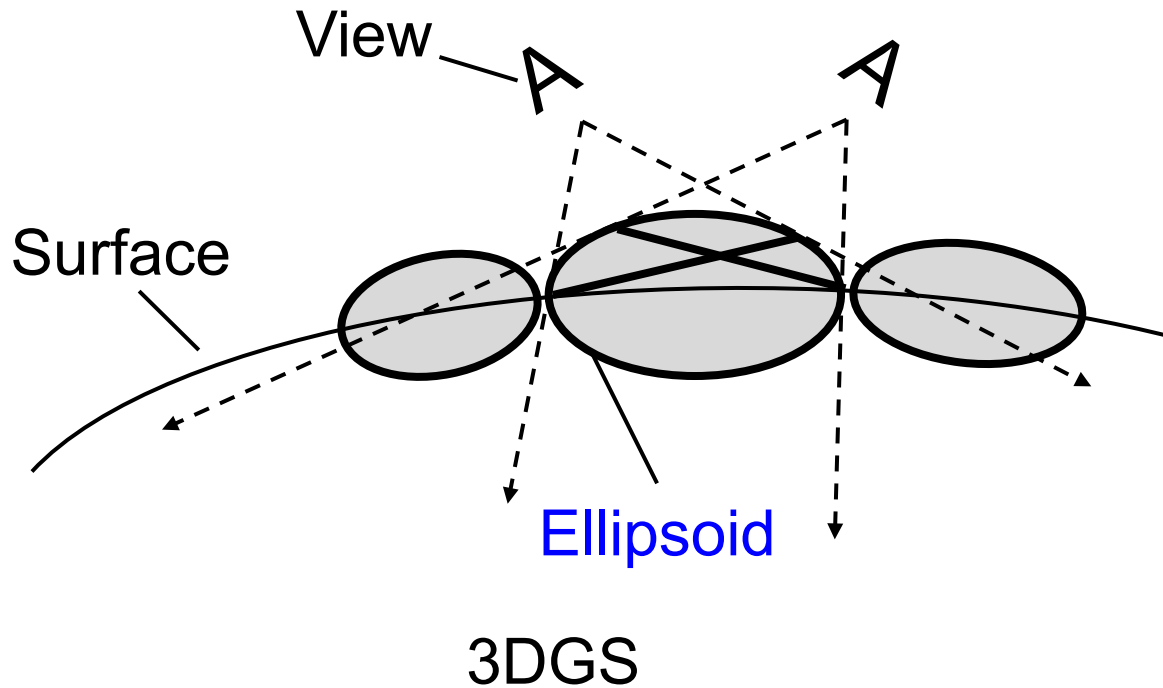
Ablation Study of Initialization (Qualitative)



- We can confirm that combining the point clouds from MVS and DUS3R enhances the reconstruction accuracy

Why 2DGS instead of 3DGS ?

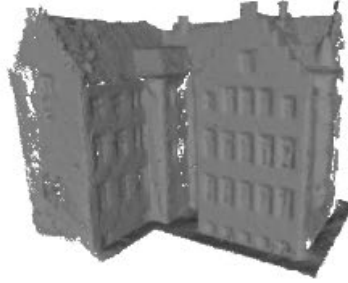
- In the performance of mesh reconstruction, 2DGS outperforms 3DGS
 - 3DGS: **Different** intersection planes in different views
 - 2DGS: **Consistent** intersection planes in different views
 - Mesh should be consistent in different views



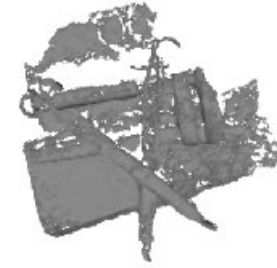
Another Reconstruction Results (Part 1)



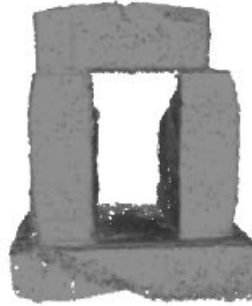
scan24



scan37



scan40



scan55



scan63



scan65



Another Reconstruction Results (Part 2)



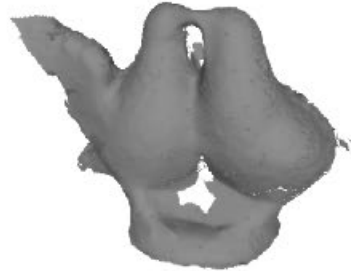
scan83



scan105



scan106



scan114



scan118

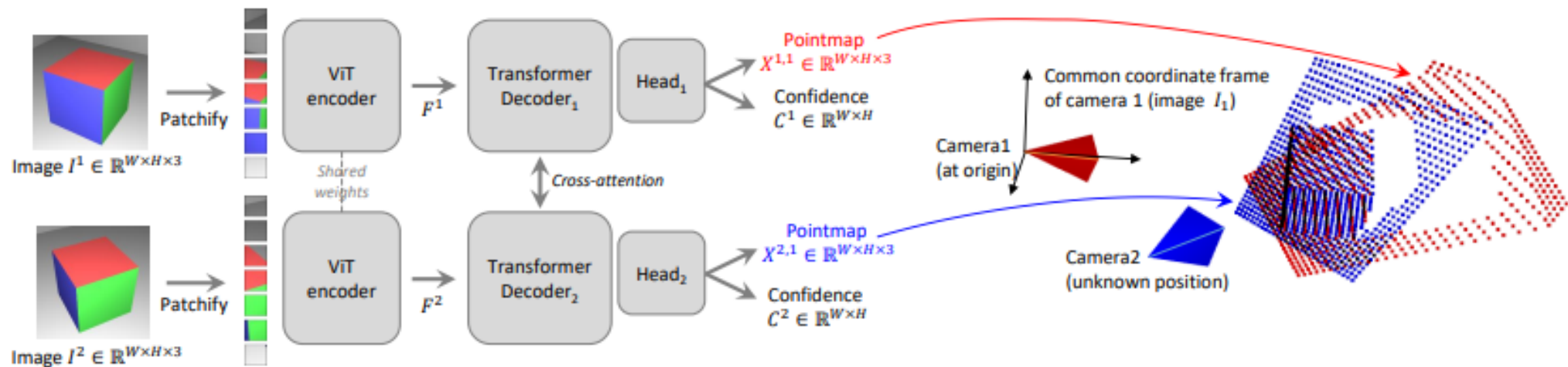


scan122



What is DUSSt3R ?

- DUSSt3R is a foundation model of stereo vision
- DUSSt3R can reconstruct a dense 3D point cloud from pointmaps



- In recent years, many methods inspired by DUSSt3R have been proposed
 - MAST3R (ECCV'24), MUST3R (CVPR'25), VGGT (CVPR'25), ...
- We think many papers inspired by VGGT will appear in the near future

UFORecon [8]

- Generalizable neural implicit surface reconstruction method based of NeRF

